Derivation of Formulas for Poisson Power Calculation

Adam Sacarny

February 9, 2021

Here I work through the Poisson power calculation formulas derived in Signorini (1991).

Consider the poisson model in which individual i's mean function is given by:

$$\lambda_i = t_i \exp\left(\beta_0 + \beta' x_i\right)$$

where t_i is exposure time (assumed independent of x_i), β_0 is a constant term, and $\beta = (\beta_1, \ldots, \beta_p)$ is a vector of coefficients. Signorini shows that the ML estimate $\hat{\beta}$ is asymptotically distributed:

$$N\left(\beta, \frac{\exp(\beta_0) M^{-1}(\beta)}{N\mu_T}\right)$$

where N is the sample size, μ_T is the mean of the exposure time , and $M(\beta)$ is derived from differentiating the moment generating function of the covariates X.

We are interested in testing the hypothesis:

$$\begin{aligned} H_N: \quad & \beta = \left(\beta_1^N, \beta_2, \ldots, \beta_p\right) \\ H_A: \quad & \beta = \left(\beta_1^A, \beta_2, \ldots, \beta_p\right) \end{aligned}$$

Let $V(\beta_1) = \left\{M^{-1}(\beta)\right\}_{22}$, the element of M corresponding to β_1. Then under H_N, $\beta_1 \sim N\left(\beta_1^N, \frac{\exp(\beta_0)V\left(\beta_1^N\right)}{N\mu_T}\right)$ and under H_A, $\beta_1 \sim N\left(\beta_1^A, \frac{\exp(\beta_0)V\left(\beta_1^A\right)}{N\mu_T}\right)$.

Thus under H_A:

$$\begin{aligned} & \Pr\left(\hat{\beta}_1 - \beta_1^N > z_\alpha \sqrt{\frac{\exp(\beta_0) V\left(\beta_1^N\right)}{N\mu_T}}\right) \\ & = \Pr\left(\frac{\beta_1^A - \hat{\beta}_1}{\sqrt{\frac{\exp(\beta_0)V\left(\beta_1^A\right)}{N\mu_T}}} < \frac{\left(\beta_1^A - \beta_1^N\right) - z_\alpha\sqrt{\frac{\exp(\beta_0)V\left(\beta_1^N\right)}{N\mu_T}}}{\sqrt{\frac{\exp(\beta_0)V\left(\beta_1^A\right)}{N\mu_T}}}\right) \\ & = \Phi\left(\frac{\left(\beta_1^A - \beta_1^N\right)\sqrt{N\mu_T \exp(-\beta_0)} - z_\alpha\sqrt{V\left(\beta_1^N\right)}}{\sqrt{V\left(\beta_1^A\right)}}\right) \end{aligned}$$

And thus we achieve power ρ when:

$$\frac{\left(\beta_1^A - \beta_1^N\right)\sqrt{N\mu_T \exp\left(-\beta_0\right)} - z_\alpha \sqrt{V\left(\beta_1^N\right)}}{\sqrt{V\left(\beta_1^A\right)}} > z_\rho$$

Rearranging shows that to achieve power ρ we must have sample size:

$$N > \frac{\exp\left(-\beta_0\right)}{\mu_T\left(\beta_1^A - \beta_1^N\right)^2}\left[z_\alpha\sqrt{V\left(\beta_1^N\right)} + z_\gamma\sqrt{V\left(\beta_1^A\right)}\right]^2$$

Then all power calculations flow from the above given $V\left(\beta_1\right)$. This is all as given in Signorini (1991) with the minor addition of allowing for a null hypothesis other than $\beta_1^N = 0$.

Suppose that there is over- or under-dispersion so that $\mathrm{Var}\left[Y_i|X_i\right] \neq \mathbb{E}\left[Y_i|X_i\right]$. If we allow that $\mathrm{Var}\left[Y_i|X_i\right] = \sigma^2\mathbb{E}\left[Y_i|X_i\right]$, the above power calculations hold by redefining $V\left(\beta_1\right) \equiv \sigma^2 V\left(\beta_1\right)$. σ^2 can be estimated from the Poisson goodness of fit tests available in Stata by running `estat gof` after a Poisson regression and dividing the fit statistic by its degrees of freedom.

In the univariate case, when X is distributed $\mathrm{Bernoulli}\left(\pi\right)$, we have $V\left(\beta_1\right) = \left(\pi e^\beta\right)^{-1} + \left(1-\pi\right)^{-1}$.

In the multivariate case, suppose X_1 is distributed $\mathrm{Bernoulli}\left(\pi\right)$ independently of $\left(X_2, \ldots, X_p\right)$, which is distributed $N\left(\mu, \Sigma\right)$, and let $\tilde{\beta} = \left(\beta_2, \ldots, \beta_p\right)$. Then we have $V\left(\beta_1\right) = \left[\left(\pi e^\beta\right)^{-1} + \left(1-\pi\right)^{-1}\right]\kappa$ where $\kappa = \exp\left(-\tilde{\beta}'\mu - \frac{1}{2}\tilde{\beta}'\Sigma\tilde{\beta}\right)$.

References

Signorini, David F. "Sample size for Poisson regression" *Biometrika* (1991)